Concept Title:	**Integrated Statistical Laboratory (ISL)**
Concept Number:	**CB-SM-94-02-E**
Concept ID:	**IT**

PART I - INFORMATION TECHNOLOGY ARCHITECTURE PLAN

1. Information Requirements

A. Introduction

The mission of the Statistical Design, Methodology, and Standards (SDMS) directorate is to provide leadership to the Census Bureau toward continuous improvement of the accuracy, cost-effectiveness, timeliness, and relevance of its products. It achieves these goals by working in collaboration with other divisions in the Census Bureau. While its agenda is shaped by the opportunities for practical improvements in Census Bureau programs, it contributes to both the theory and practice of demographic and economic measurement through statistical, methodological, and technological research. It consults broadly within and outside the Census Bureau to enhance the efficiency of applying research findings to the Census Bureau's decision making process.

B. Information Flow

Within the SDMS directorate, the mission of the Integrated Statistical Laboratory (ISL) is to support the Statistical Research Division (SRD) which includes the Center for Survey Methods Research (CSMR).

SRD has a broad research agenda in the following areas:

- Support for HPC Research
- Record Linkage
- Automated Coding
- Small Area Estimation
- Variance Estimation
- Survey Methods Research
- Questionnaire Design
- Computer-Assisted Survey Information Collection (CASIC) Technology
- Time Series
- Undercount Research
- Quality Control
- Survey Sampling
- Editing and Imputation
- Cognitive Methods
- Ethnographic Research

Research and prototyping activities of new CASIC technologies are being accomplished within the framework of the ISL network. These projects include Touchtone Data Entry (TDE), Voice Recognition Entry (VRE), Large Vocabulary Recognition Entry (LVRE), and FAX reporting and transmission of forms.

The mission of the ISL, serving the SDMS directorate, is to provide a statistical research computing environment that provides the power and flexibility to handle multiple operating systems, many users, large files and databases, a large variety of statistical, analytical and graphical packages and libraries, highly compute-intensive customized programs, and the network communications to permit researchers to share information easily and accurately.

The Census Bureau component involved in carrying out this concept is SRD which includes CSMR. These offices are under the administrative control of the Associate Director for Statistical Design, Methodology, and Standards. The ISL computer facility resources are available to all Census Bureau directorates, and joint research projects across directorates are strongly encouraged.
Projects which cross directorates require the ability to access and transfer files over the network. This requires that support and cooperation with telecommunications is a necessity.

2. Planned Processing and Telecommunications Architecture

A. Current Architecture

The current ISL configuration is a mixed network using both the older coaxial cable plant and the newer unshielded twisted pair (UTP) environment. The migration to the new UTP environment has been slowed by hardware, software, and network considerations. There are 2 primary UNIX compute servers (a Hewlett-Packard (HP) 825 and a SUN SparcCenter 2000), 1 primary file server (a Solbourne 5/605), 22 UNIX-based workstations, 20 X-terminals, 3 NOVELL PC-LAN file servers, and 102 Intel-based PCs . Other devices, including local printers and plotters, and network controlling devices, including repeaters, gateways, routers, and filters number approximately 70. These devices support 100 staff in SRD and CSMR, as well as users in other divisions who collaborate on research projects.

Most general purpose computation in the ISL is done on the SparcCenter and UNIX workstations. The SOLBOURNE is a 5 CPU, 128MB RAM multi-processor machine which is binary compatible with SUN computers which now serves primarily as a file server and backup to the SparcCenter. Each CPU provides about 20 million instructions per second (MIPS), with multi-processor throughput equivalent of about 90 MIPS. The SOLBOURNE has 18 gigabytes of disk storage. Communications software on the SOLBOURNE allows file transfer and login capabilities to the Census Bureau's Digital computers as well as any other UNIX machines on the Census Bureau's network. The SparcCenter 2000 has 8 CPUs and 66 GB of storage.

The 22 Unix workstations are SUN processors of various models and age (except for one HP 730), that vary substantially in performance, disk capacity and display capabilities. These workstations provide CPU and local storage dedicated to individual projects, as well as a high resolution graphical display.

The 20 X-terminals are a substantially less expensive alternative to workstations, yet provide a highly responsive X-window environment that communicates easily to any VAX or UNIX machine on the network.

The ISL UNIX platforms provide programming and software development tools, data analysis software, graphical packages for statistics and presentations, and advanced mathematical and statistical software. These software tools are critical parts of nearly every research project, and it is essential for effective research that these be highly capable.

The HP server is a small 8 MIP processor dedicated to supporting the work of the Industry and Occupation coding project, which is joint between the Housing and Household Economic Statistics Division (HHES) and SRD. This device is scheduled to be decommissioned with its functions being served by an upgraded HP730 workstation.

The PCs and NOVELL servers provide basic office automation applications. The NOVELL servers are single points of failure, and a server loss would disrupt the office work of the affected division. For this reason the second NETWARE server in SRD is a dedicated backup server synchronized to the primary server. With the merger of SRD and CSMR into one division and the physical realignment of the divisions in FB-4, the requirement for a single merged server is necessary.

B. Alternatives

The ISL represents a model of the open systems environment toward which the Bureau is aiming. The challenge is to integrate the PC and workstation environments with a solid network to allow any program to meet its IT needs in as efficient a way as possible.

C. Proposed Architecture

Equipment purchased a few years ago is no longer meeting the computational requirements of the researchers. As an example, a SAS application on a 386 class PC will execute but the total time required to complete the job is so long that executing the job on this platform is essentially impossible. Similar bottlenecks will occur as the needs of SDMS researchers increase while the computing capability stagnates.

Although the combined compute power of all the ISL machines is impressive when compared to what was available just a few years ago, the nature of research is such that there is always a

demand for more. Researchers require the use of larger and more complete data files rather than extracts of files to which they were limited only a few years ago.

Obsolescence of the current stock of ISL equipment is an increasing factor in the performance of the overall system. Communication problems are caused by an antiquated cable plant, disparate speeds among the devices on the network. Lack of software support for aging operating systems and architectures constrains the useability of both old and new software. Approximately 30 percent of the PCs are a 386 class or below making it difficult to service this resource and to standardize on a robust graphical user interface. The ISL intends to participate in PC replacement strategy promulgated in the Personal Computer Management Acquisition Plan (PCMAP).

Fortunately, it is not necessary to make an increasing capital investment in new computer equipment to meet this demand. Since the cost of computing continues to decline, the same rate of investment produces the necessary increased computing capacity.

This concept is addressed primarily to the continuing modernization of existing computing resources in order to maintain a leading edge statistical computing system for research. Failure to modernize would lead to reduced research capability, and less opportunity to improve methodology for Census Bureau surveys and censuses.

The ISL equipment represents a wide range of useability from obsolete to nearly state-of-the-art. In order to continue to function as a reliable network, older, slower devices, especially the 286 and 386 PCs and older workstations, the SUN ELCs and IPCs must be upgraded or replaced. Otherwise, interruptions and decreased reliability of backup functions will occur. One essential improvement is the replacement of the 10Base2 cabling system with 10BaseT (Twisted pair) as a minimum to providing the bandwidth need in the future. Other technologies such as ATM and direct FDDI connections should be seriously considered, especially as part of a Census Bureau-wide upgrade in this area.

As the Bureau moves toward an Open Systems environment, the ISL will fit more easily into the Bureau structure as the ISL was designed as an open system. The enhanced Bureau network strategy should allow for all devices on the network to be shared and to maximize the use of dedicated devices for computing, file sharing, and backup.

D. Benefits

The primary benefit of the ISL is a strengthened research effort for the Census Bureau. A strong research program focuses on the future, seeking continuous improvements in the accuracy, cost-effectiveness, timeliness, and relevance of Census Bureau products.

An additional benefit of the ISL is to create the environment where solutions to a research problem can be investigated in as many ways as possible rather than being restricted to solutions

constrained by hardware and/or software inadequacies. In essence the ISL functions as one of the test beds for the open-system environment envisioned for the Census Bureau.

The current ISL configuration allows for the continuing upgrade of current products with the new technology in the marketplace without sacrificing older resources to complete obsolescence. The goal is to keep older products useful while integrating better resources to projects while minimizing the costs associated with switching to new technologies. Another important feature here is the judicious use of software portable across platforms and the use of software tools that optimize applications easily across platforms.

3. Security

Privacy and security will be maintained by provisions of Title 13, United States Code, the Census Administrative Manual, and applicable sections of the Department of Commerce (DOC) Handbook of Security Regulations and Procedures. Procedures to ensure privacy and security for existing systems will also apply to this new equipment.

Concept Title:	**Integrated Statistical Laboratory (ISL)**
Concept Number:	**CB-SM-94-02-E**
Concept ID:	**IT**

PART II - ANNUAL PLAN

1. Architecture Status

The ISL was designed to be easily upgradeable in concurrence with the Bureau's Open Systems philosophy and will remain a viable and important Bureau-wide resource for the next 5 years. The ISL, using SRD funds, has provided for both hardware and software maintenance contracts on the necessary equipment. This will continue to the case unless the Bureau institutes a corporate strategy to provide these services for all equipment of this type.

2. IT Objectives

The basic objective for FY 1997 is to maximize network performance through a combination of both hardware and software upgrades.

Specific upgrade proposals are:

- Complete the conversion to UTP and minimizing (with TCO's assistance) the level of traffic required to traverse the Census backbone network.
- Replace the Solbourne as a file server.
- Upgrade all remaining 386 PC's to the current Bureau standard using the PCMAP guidelines.
- Standardize on the X-window environment using a combination of PC's with the necessary software and X-terminals.
- Establish the primary office automation products of electronic mail and word processing in the UNIX environment.
- Upgrade and standardize all servers and PC's to the same operating system (Solaris 5.3, DOS 6.22 and Windows for Workgroups 3.11). A transition to Windows 95 for the PC environment will be in accordance with the policy of the Bureau's Telecommunications Office.

There are no planned major hardware and software upgrades or purchases for FY 1998.

3. Status

During 1996 the ISL continued to upgrade its resources via replacement of older PCs, upgrades to existing servers and improved backup resources including faster and more reliable tape drives and larger capacity tapes.

4. Implementation Schedule

Milestone	Actual Start	Planned Completion
Integration of twisted pair topology to desktop	10/95	04/96
Acquire and install ISL upgrade hardware and software	10/95	09/96